IP over ATM

Classical IP over ATM, MARS, MCS,
NHRP, LANE, MPOA

Agenda

- <u>Introduction</u>
- Encapsulation
- Classical IP and ARP over ATM (ARP Server)
- Multicast Support (MARS, MCS)
- LANE
- Next Hop Resolution Protocol (NHRP)
- MPOA

IP over ATM

- **ATM is connection-oriented**
  - Assumes connection-oriented applications
- **IP is connection-less**
  - Assumes connection-less network
- **Significant mismatch**
- **How to solve the problem**
  - Interface layer between IP and ATM
    - connection-less behavior towards IP
    - connection-oriented behavior towards ATM
  - Several methods
    - IETF approach
    - ATM Forum approach

IP over ATM : Solving the Problem I.

- **Encapsulation**
  - If IP is transmitted over any type of media, an appropriate encapsulation has to be defined
  - IETF RFC 2684 (former 1483)
  - IETF RFC 2364 (PPPoA)
- **Address resolution**
  - IP and ATM addresses are different
  - IETF RFC 2225 (former 1577) (ATM-ARP Server)
  - *ATM Forum LAN Emulation*
- **Transmission unit**
  - A maximum transmission unit has to be defined
  - IETF RFC 2225 (former 1626)

IP over ATM: Solving the Problem II.

- **Connection-setup**
  - ATM requires connections
  - IETF RFC 1755 based on UNI 3.1
  - IETF RFC 2331 based on UNI 4.0
- **Broadcast and multicast support**
  - ATM does not inherently support broadcasts
  - IETF MARS (RFC 2022, Multicast Address Resolution Server)
  - IETF MCS (RFC 2149, Multicast Server)
  - *ATM Forum LAN Emulation*

IP over ATM: Solving the Problem III.

- **Path determination**
  - ATM has it´s own routing
  - ATM is treated as NBMA
    - Non Broadcast Multi Access
  - Hop-by-hop IP routing not appropriate for ATM networks
  - IETF Next Hop Resolution Protocol (NHRP, RFC 2332)
  - IETF Multi Protocol Label Switching (MPLS, RFC 3031)
  - *ATM Forum Multi Protocol over ATM (MPOA)*
  - *ATM Forum Integrated PNNI*

Agenda

- Introduction
- Encapsulation
- Classical IP and ARP over ATM (ARP Server)
- MTU and Signaling
- Multicast Support (MARS, MCS)
- LANE
- Next Hop Resolution Protocol (NHRP)
- MPOA

Encapsulation

- A standards based method is needed to identify the protocol within a carried PDU
  - Comparable to ETHERTYPE, PPP, LLC, RFC 1490, etc.
- Two methods
  - LLC encapsulation
    - Using LLC header or SNAP
  - VC multiplexing
    - N protocols means N channels
  - Defined in RFC 2684
- PPPoA (PPP over ATM
  - Defined in RFC 2364
- AAL5 as adaptation procedure

LLC Encapsulation

- **Is needed if several protocols are carried over the same virtual circuit**
  - Called multi protocol encapsulation over ATM
- **Payload field must contain information to identify the protocol of the routed or bridged PDU**
  - This information is encoded in an LLC header placed in front of the PDU
  - Can be a simple LLC Type 1 header or a SNAP header

SNAP

AA_1	AA_1	03_1	OUI_3	$Protocol_2$	$Data_{>=0}$

802.2 LLC Type 1

$DSAP_1$	$SSAP_1$	$CTRL_1$	$NLPID_{1or2}$	$DATA_{>=0}$

802.2 LLC Type 1 for PPPoA

0xFE	0xFE	0x03	0xCF	$DATA_{>=0}$

LLC Encapsulation

- **Different specifications for**
  - ISO routed PDUs
    - Use standard LLC header (FE-FE-03)
    - Standard LLC Header followed by NLPID
    - NLPID = 0xCF is used for PPPoA
  - NON-ISO routed PDUs
    - Use SNAP header (AA-AA-03-00-00-00-ETHERTYPE)
    - That is used for IP
  - Bridge PDUs
    - Use SNAP header plus IEEE 802.1organizational code
    - Type of bridged network is specified by two octet PID
    - PID also specifies whether original FCS is maintained or not
    - Padding is used to align the user information field to a four octet boundary

LLC Encapsulation

- **Summary of encapsulations**

Format	FCS	LLC	SNAP OUI	SNAP PID	PAD / FC	FCS
Routed ISO PDUs	na	FE-FE-03	na -> NLPID	na	na	na
Routed Non-ISO PDUs	na	AA-AA-03	00-00-00	ETHERTYPE	na	na
Routed IP PDUs	na	AA-AA-03	00-00-00	08-00	na	na
Bridged 802.3	yes	AA-AA-03	00-80-C2	00-01	00-00	CRC
Bridged 802.3	no	AA-AA-03	00-80-C2	00-07	00-00	na
Bridged 802.4	yes	AA-AA-03	00-80-C2	00-02	00-00-00-FC	CRC
Bridged 802.4	no	AA-AA-03	00-80-C2	00-06	00-00-00-FC	na
Bridged 802.5	yes	AA-AA-03	00-80-C2	00-03	00-00-XX-FC	CRC
Bridged 802.5	no	AA-AA-03	00-80-C2	00-09	00-00-XX-FC	na
Bridged 802.6	na	AA-AA-03	00-80-C2	00-0B	Note	na
Bridged FDDI	yes	AA-AA-03	00-80-C2	00-04	00-00-00-FC	CRC
Bridged FDDI	no	AA-AA-03	00-80-C2	00-0A	00-00-00-FC	na
Bridged PDU	na	AA-AA-03	00-80-C2	00-0E	na	na

FCS .. is FCS maintained
FC .. Frame Control
Note: with 802.6, this field contains the Common PDU Header (Reserved, BEtag, BAsize)

VC Multiplexing

- **Multiple protocols are transmitted using multiple VCs**
  - Protocol is identified by the VC itself
- **Only one protocol may be carried over a single VC**
- **No need for multiplexing information**
  - Minimal overhead
- **Two formats**
  - Routed protocols
    - Are transmitted as is
  - Bridged protocols
    - Are transmitted as described under LLC encapsulation

Agenda

- Introduction
- Encapsulation
- Classical IP and ARP over ATM (ARP Server)
- Multicast Support (MARS, MCS)
- LANE
- Next Hop Resolution Protocol (NHRP)
- MPOA

Classical IP and ARP over ATM

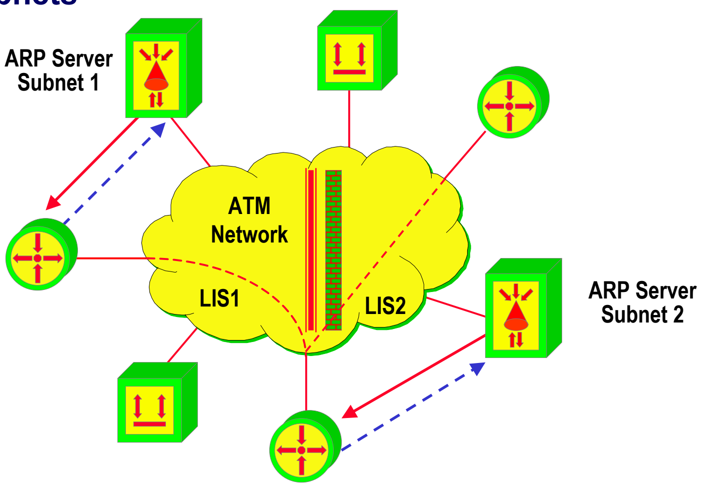
- **Address resolution required**
  - Link layer address has no relationship with IP address
    - No relation between MAC address and IP address
  - ARP used with classical LANs
    - Requires broadcast support
- **ATM does not support broadcasts**
  - ARP had to be modified
  - Uses ARP server instead of broadcasting
- **Defined in RFC 2225 (former 1577)**

Classical IP and ARP over ATM

- **RFC 2225 does not solve the broadcast issue**
  - Not a solution for interconnecting routers
  - Routing updates require broadcast support
    - Permanent VC has to be established between routers
- **One ARP-server per IP subnet**
  - Multiple ARP servers needed, if ATM cloud is separated into multiple subnets
  - Communication between subnets requires router
    - Performance problem
- **Specifies one single ARP server**
  - Single point of failure, redundancy is up to implementers

RFC 2225 Operation

- **multiple hops for communication between logical IP subnets**



RFC 2225 Operation: Registration

- Has to be initiated by the client
  - Requires configuration of the ARP server address
- Client connects to ARP server using a standard virtual circuit (SVC)
- ARP server issues IN_ARP_request (inverse ARP)
- Client answers with IN_ARP_reply
  - Contains ATM as well as protocol address of client
- ARP server places information in its cache
  - Builds address resolution table

RFC 2225 Operation: Connection Setup

- Client sends ARP_request to server
- Server looks up desired destination in its table
  - If there is an entry, server will respond with an ARP_reply
  - If there is no entry, server will respond with an ARP_NAK
- Client uses information from ARP_reply to setup a call to target ATM address

RFC 2225 Operation: Caching

- Server as well as client maintain an ARP cache
  - Server aging time = 20 minutes, client aging time = 15 minutes
  - Server refreshes cache by looking at ARP_requests
  - Server refreshes cache by placing IN_ARP_requests on every open VC on a periodic basis
- If client has no open VC to server, client needs to re-register every 20 minutes
  - Client refreshes cache by placing a new request to the ARP server (if SVCs are used)
  - Client refreshes cache by placing an IN_ARP_request on any open VC (if PVCs are used)

 IP over ATM, v4.2 19

IP MTU over ATM

- Defined in RFC 2225 (former 1626)
- Shall be reasonably large in order to avoid fragmentation
- Defined to be 9180 octets as for SMDS (RFC 1209)
- RFC 2225 defines furthermore the use of Path MTU discovery as defined in RFC 1191
  - Router implementations must implement RFC 1191
  - Host implementations should implement RFC 1191

 IP over ATM, v4.2 20

Signaling for IP over ATM

- **Outlined in RFC 2331 (former 1755)**
- **Presents implementation guidelines**
  - For the support of UNI 3.1 signaling (RFC 1755)
  - For the support of UNI 4.0 signaling (RFC 2331)
  - Details the use of VCs in conjunction with RFC 1577
  - Details signaling, traffic descriptors, etc.
- **Specifies how and when to tear down a VC**
  - Suggests a minimum holding time of 60 seconds
  - Specifies a configurable inactivity timer to clear idle connections
    - MUST be implemented at the public UNI
    - SHOULD be implemented at the private UNI

 IP over ATM, v4.2 21

Agenda

- **Introduction**
- **Encapsulation**
- **Classical IP and ARP over ATM (ARP Server)**
- **Multicast Support (MARS, MCS)**
- **LANE**
- **Next Hop Resolution Protocol (NHRP)**
- **MPOA**

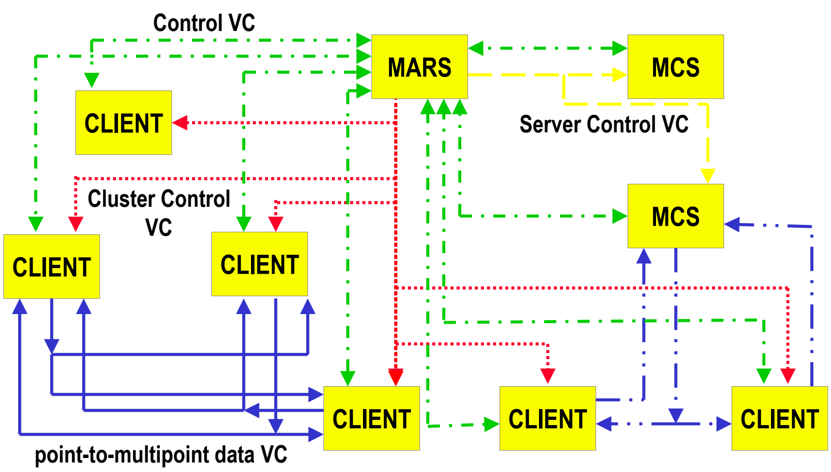
 IP over ATM, v4.2 22

Multicast Support

- **Two issues**
  - Resolving ATM destination addresses for a given layer 3 multicast group
  - Sending data to members of multicast group
- **Multicast Address Resolution Server (MARS)**
  - Solves the address resolution problem
  - Defined in RFC 2022
- **Sending data to members of a group**
  - Direct distribution using a mesh of point-to-multipoint circuits
  - Multicast Server (MCS), defined in RFC 2149

 IP over ATM, v4.2 23

MARS/MCS Architecture



 IP over ATM, v4.2 24

Multicast Support

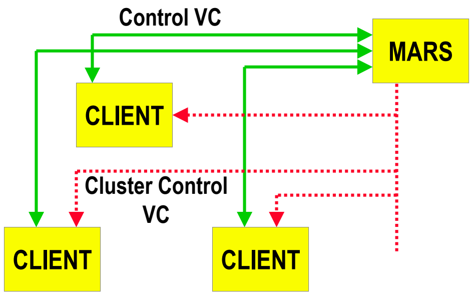
- **Layer 3 multicasting usually supported by sending data to a layer 2 multicast address**
  - ATM does not support layer 2 multicasting
  - Furthermore, there is no relation between layer 3 multicast addresses and ATM addresses
- **Multicast Address Resolution Server**
  - Enhancement of ATM-ARP server (RFC 2225)
  - Maintains a list of ATM addresses per layer 3 multicast address
  - MARS operation is protocol independent
    - Does support protocol identification
  - Broadcast support not explicitly defined, but possible

MARS Operation

- **Clients are configured with ATM address of MARS**
  - At least, one MARS needs to be configured
- **MARS redundancy**
  - More than one MARS may be configured
  - If first MARS does not respond or fails, client connects to alternate MARS
- **Two modes of operation**
  - Direct connectivity between members of a multicast group
    - mesh of point-to-multipoint VCs between clients
  - Multicast Server acting as a proxy for clients

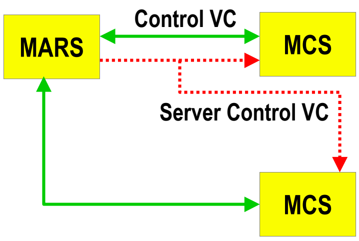
MARS Operation

- **Initialization phase**
  - Client opens point-to-point VC to MARS
  - Registers itself as a multicast capable device
    - MARS_JOIN message, includes client ATM address
  - MARS adds client to its Cluster Control VC
    - Client receives 16 bit Cluster Member Identification (CMI)
    - Point-to-multipoint VC



MARS Operation with MCS

- **initialization phase**
  - MCS opens point-to-point VC to MARS
  - Registers itself as a multicast server
    - MARS_MSERV message, includes ATM address
  - MARS adds MCS to its Server Control VC

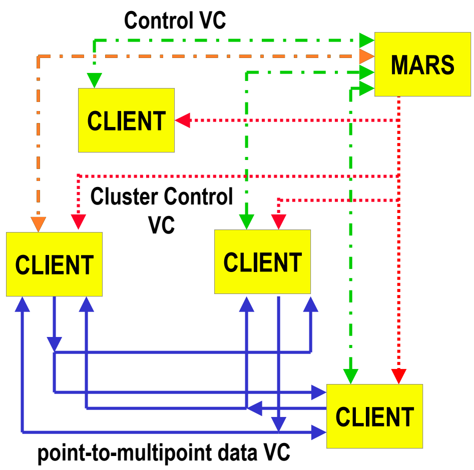


MARS Operation

- **Joining or leaving a multicast group**
  - Client uses point-to-point VC to MARS
  - Sends MARS_JOIN, indicating layer 3 multicast address
  - May send MARS_LEAVE to leave a multicast group
  - Information is passed to other clients using Cluster Control VC
- **MCS operation**
  - Similar to client
  - Uses MARS_JOIN, indicating the support of a layer 3 multicast address

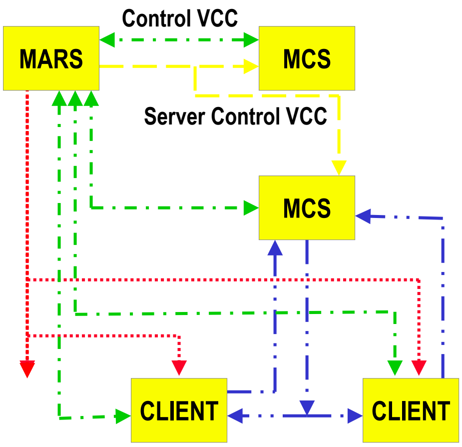
Multicasting without MCS

- **Client queries MARS**
  - Sends MARS_REQUEST, indicating layer 3 MC group
  - Receives MARS_MULTI, containing list of ATM addresses
    - ATM addresses of other clients (group members)
  - Opens point-to-point VC to members of group
- **All other clients perform the same operation**



Multicasting with MCS

- MARS maintains two tables
  - Client table, MCS table
- MCS queries MARS
  - Receives list of clients
  - Establishes pt-to-multipoint circuit to clients
- Client queries MARS
  - receives list of MC servers
    - multiple servers supported
  - establishes circuit to MCS
- MCS passes data traffic received from clients to other clients



Updating Group Membership

- Without MCS
  - MARS passes MARS_JOIN and MARS_LEAVE messages to other clients
    - Using the Cluster Control VC
    - Allows clients to keep track of group joins and group leaves
- With MCS
  - MARS passes MARS_JOIN and MARS_LEAVE messages to MCS
    - Special messages MARS_SJOIN and MARS_SLEAVE
    - Allows MCS to keep track of group joins and group leaves
  - Note: MARS_JOIN and MARS_LEAVE messages are NOT distributed to clients using Cluster Control VC

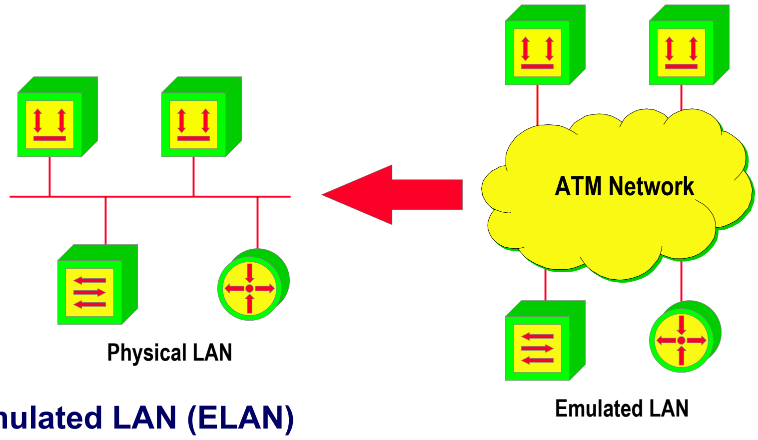
Agenda

- Introduction
- Encapsulation
- Classical IP and ARP over ATM (ARP Server)
- Multicast Support (MARS, MCS)
- LANE
- Next Hop Resolution Protocol (NHRP)
- MPOA

Overview

- Generalized MAC bridging approach
  - Works for all network protocols
  - Emulates MAC addresses and broadcast support
- Requires several server functions
  - Broadcast support
  - Mapping between ATM and MAC addresses
  - Configuration information
- Different for Ethernet and Token Ring
  - Different bridging methods

Overview



- **Emulated LAN (ELAN)**
  - Equivalent to one legacy LAN segment
  - Provides necessary connectionless service and multicasting across the ATM network

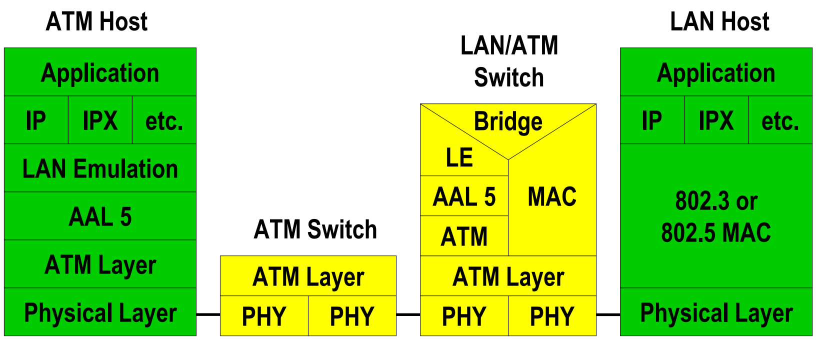
Overview

- **Ethernet LAN**
  - Any-to-all communication
    - Broadcast media network
  - Connectionless network
  - Direct UC/MC/BC support
    - Part of the protocol
  - Address mapping
    - Network ARP resolution
- **ATM LAN**
  - Point-to-point communication
    - Needs server support
  - Connection-oriented network
    - Needs signaling support
  - No direct UC/MC/BC support
    - Needs server support
  - Address mapping
    - MAC to NSAP resolution

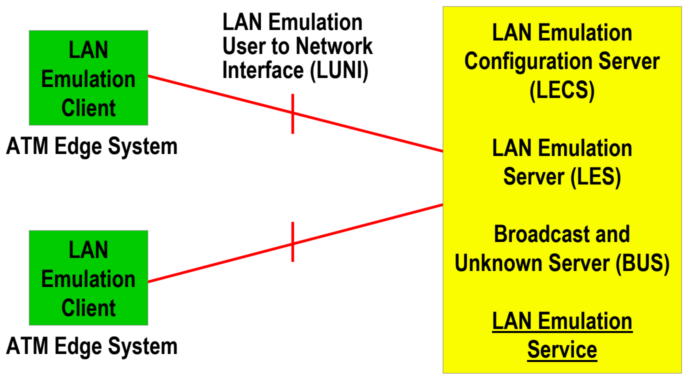
Overview

- **LANE protocol defines operation of a single emulated LAN (ELAN)**
- **Multiple ELANs may coexist simultaneously on a single ATM network**
- **A single ELAN emulates either an Ethernet or Token Ring network**
- **Each ELAN consists of a single broadcast domain**
  - Represents a physical Ethernet or Token Ring segment
- **LANE uses RFC 1483 encapsulation**
- **LANE is not part of ATM itself**
  - it is a function which uses ATM

LAN Emulation Protocol Stack

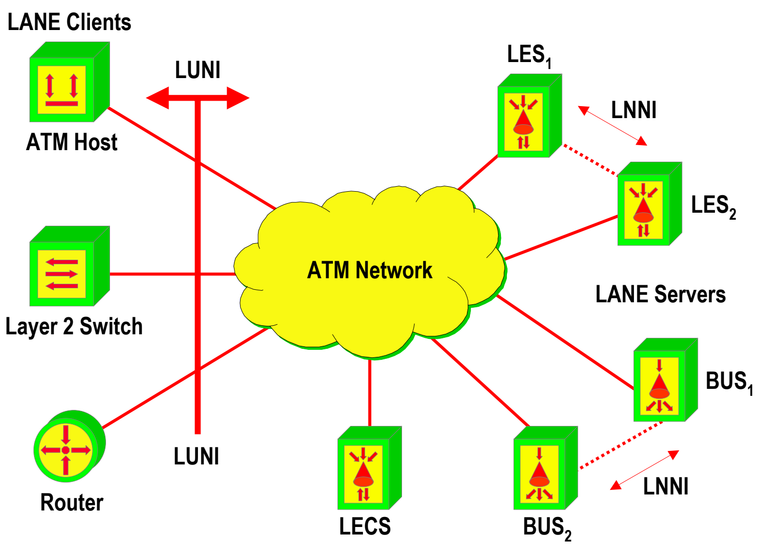


LAN Emulation Components



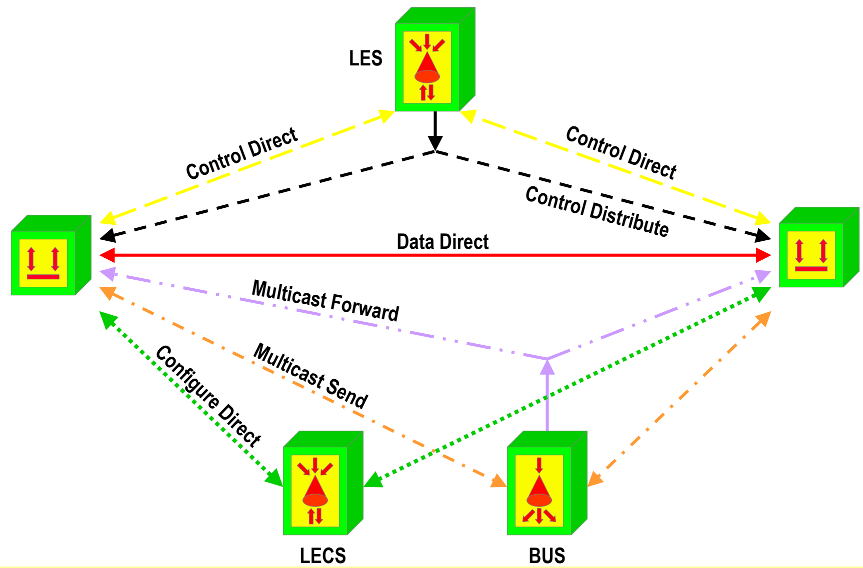
- **LECS** stores information about emulated LANs: name, type, frame size, LES ATM address, etc.
- **LES** stores information about registered MAC/ATM addresses of LAN Emulation clients
- **BUS** forwards unknown and multicast data to LAN Emulation Clients

LANE Interfaces



Note: The Phase 1 LANE spec only specifies the LUNI interface

Connections



LANE VCs

- **Configure Direct VC (LECS)**
  - used by the LEC to obtain the LES address for the ELAN it wishes to join
- **Control Direct VC (LES)**
  - used by the LEC to join the ELAN and for LE-ARP requests
- **Control Distribute VC (LES)**
  - used by the LES to forward unresolved LE-ARP requests
- **Multicast Send VC (BUS)**
  - used by the LEC to send broadcast and unresolved unicast traffic

LANE VCs

- **Multicast Forward VC (BUS)**
  – used by the BUS to forward broadcast and unresolved unicast traffic received on the Multicast Forward VCC to the remaining LECs
- **Data Direct VC (LEC-LEC)**
  – used to transfer end-to-end bulk of data traffic

LAN Emulation Client - LEC

- **Entity located in end systems**
  – Performs address resolution and data forwarding
- **Provides emulated interface for higher layers**
  – Ethernet / IEEE 802.3 or IEEE 802.5
- **Uses LUNI to communicate with other entities**
- **LEC identifier (LECID)**
  – Is contained in the LANE frame header
  – Indicates the ID of the LANE client
  – The LECID is unique for every LANE client
- **Each LEC can be only a member of a single ELAN at any time**

LAN Emulation Client - LEC

- Can be an ATM connected end station, server, router, LAN switch or bridge
- Identified by an unique ATM address
- Multiple LEC instances per physical interface have all unique ATM addresses

LAN Emulation Server - LES

- Implements the control functions for an ELAN
- Enables a LEC to register and join an ELAN
  - Clients may register the LAN destinations they represent
- Resolves address resolution issues
  - MAC to ATM address mapping
  - Clients query the LES when they wish to resolve the MAC address and / or route descriptor to an ATM address
- There is at least one active LES per ELAN
- The LES is dedicated to a single type of LANE
- The LES is identified by an unique ATM address

LAN Emulation Configuration Server - LECS

- **Implements the assignment of individual LE clients to different emulated LANs**
  - Tells a client the ATM address of the corresponding LES
  - Allows the LEC to automatically configure
  - The LECS is identified by an unique ATM address
- **Several options for a LEC to connect to the LECS**
- **Logically one LECS per administrative domain**
  - Serves all ELANs within that domain
  - Contains configuration information for all ELANs
  - Allows central configuration and administration of multiple ELANs in an ATM network

 IP over ATM, v4.2 47

Broadcast and Unknown Server - BUS

- **Handles data sent by a client to the broadcast MAC address 'FFFF.FFFF.FFFF'**
- **Handles all multicast traffic**
- **Relays broadcast or multicast packets via a point-to-multipoint VC to all LECs of a particular ELAN**
  - The source LEC uses the LECID in the frame to filter its own frames
- **Handles initial unicast frames sent by the LEC**
  - Before target ATM address has been resolved
  - Before VC to target ATM address has been established
  - Flooding Feature of transparent Bridging

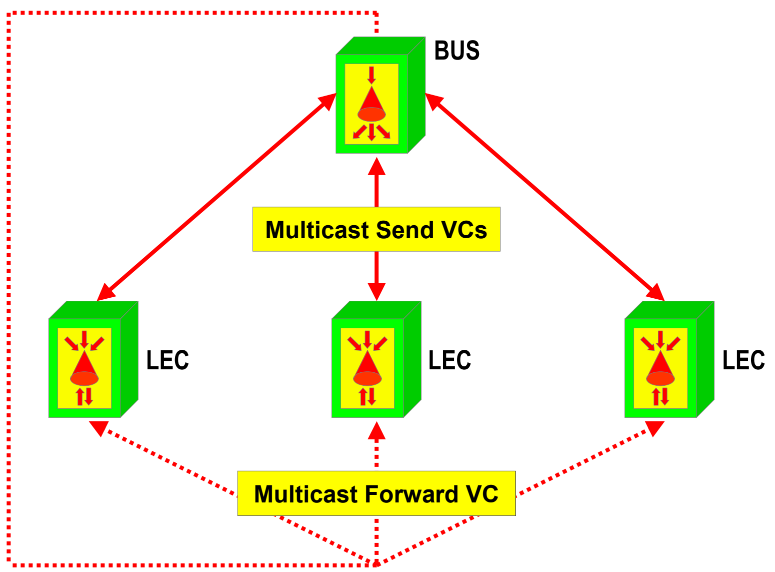
 IP over ATM, v4.2 48

Broadcast and Unknown Server - BUS

- **Transmission is done in a FIFO manner**
  - One frame at the time
- **If the BUS receives two frames simultaneously, it buffers one frame while it sends the other frame**
  - Works in a store-and-forward manner
  - An AAL 5 PDU has to be completely received before the frame can be forwarded
- **Serialization prevents intermixing of cells from different frames going to the LECs**
  - AAL 5 used with LANE allows reassembly of only one frame at a time on a single VC

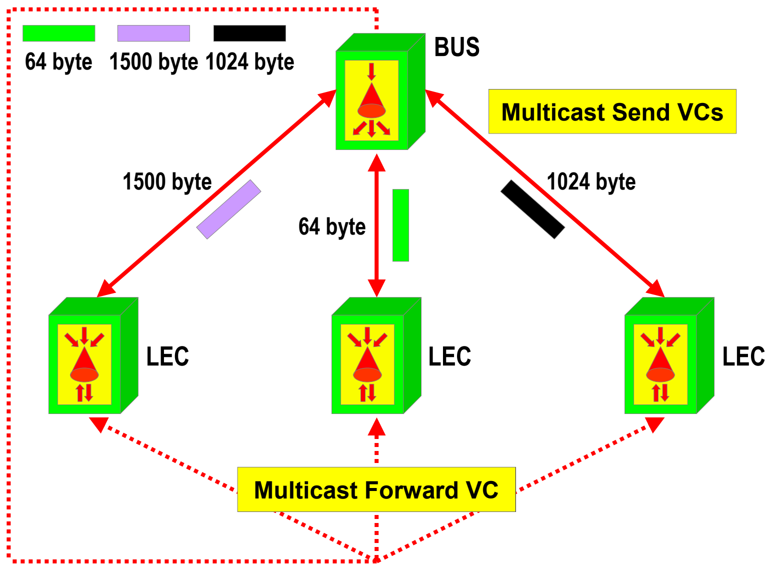
 IP over ATM, v4.2 49

BUS Traffic Serialization



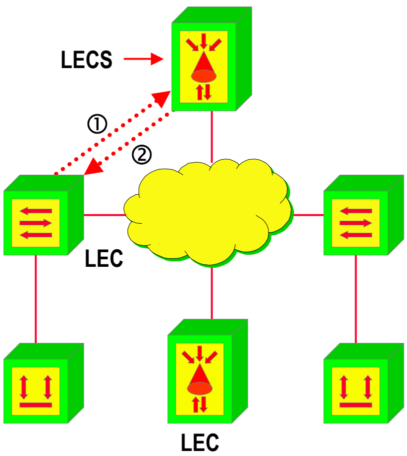
 IP over ATM, v4.2 50

BUS Traffic Serialization



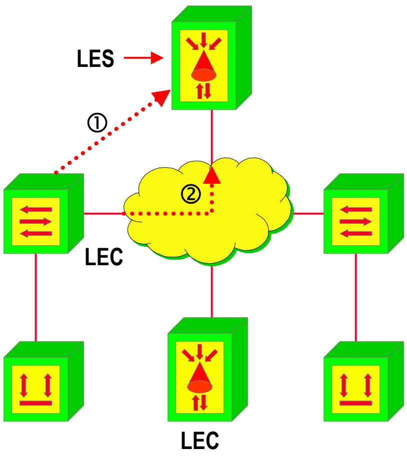
LAN Emulation - Initialization

- **Client needs to contact LECS to learn ATM address of LES for the specific ELAN**
- **May use different methods**
  - ILMI
  - configured ATM address
  - well known VPI/VCI
    - 0/17
  - well known ATM address
- **Connects to LECS**
- **Obtains ATM address of LES**



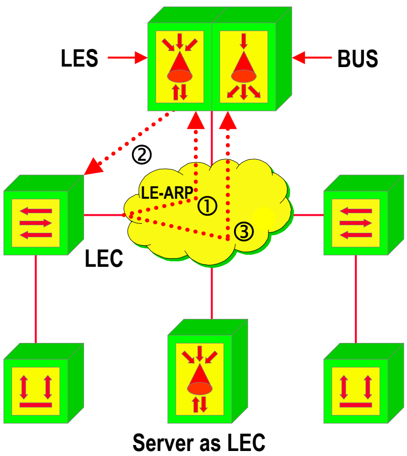
LAN Emulation - Registration

- **Client connects to LES**
  - Registers its MAC address
  - Registers a list of individual MAC addresses the client represents (i.e. LAN/ATM switch)
  - Registers a list of Source Route descriptors (segment / bridge pairs) the client represents, Token Ring only
  - Registers as a proxy - will take part of the resolution process for finding non-registered, non-ATM attached devices (must for a LAN/ATM switch)
- **LES adds LEC into its database**



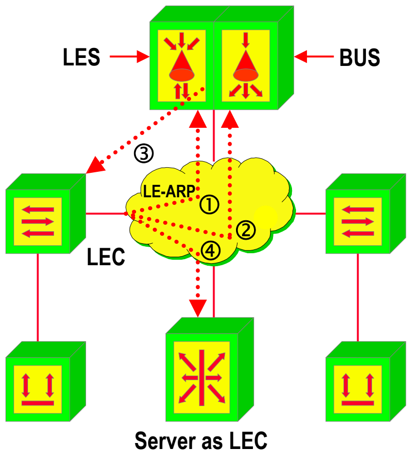
LAN Emulation - Address Resolution

- **Client resolves address of BUS**
  - Sends an ARP request to LES asking for MAC address FFFFFFFF
  - Connects to BUS
- **Client sets up a multicast send VC to the BUS**
- **BUS adds client to its multicast connection**
  - BUS sets up a multicast forward VC to the client



LAN Emulation - Data Transfer

- **Client starts resolving destination MAC address with LES**
  - At the same time, the client may start sending frames to the BUS
- **LES replies with requested information (destination ATM address)**
- **Client connects directly to the destination**
  - Client flushes bus connection
  - Client starts transmitting frames directly



Some Details

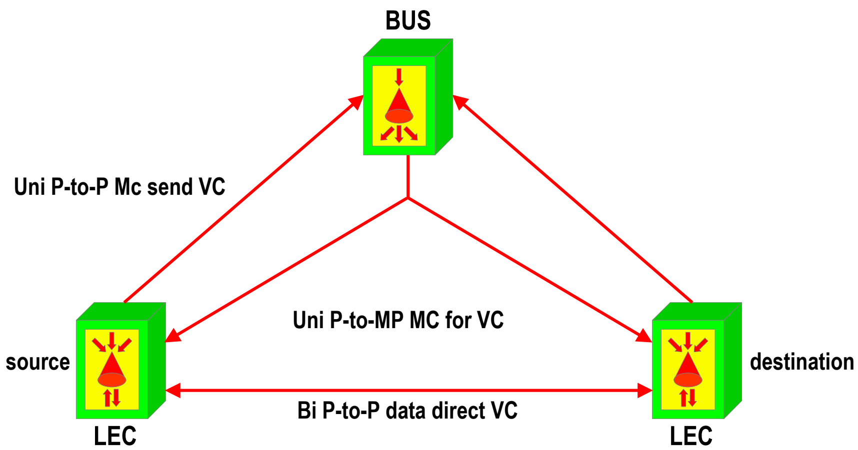
- **ELAN name**
  - SNMPv2 display string (max. 32 characters)
- **LEC Identifier (LECID)**
  - First two bytes of the LE data frame header define membership in the ELAN via the LECID
  - Assigned to that particular LEC by the LES
  - Used to identify and filter own frames by comparing the LECID field to its own LECID value
- **LECs maintain an ARP cache**
  - Contains all LE ARP replies
    - aging time is typically 5 minutes, minimum of 10 seconds

LANE Flush Protocol

- **Ensures the correct order of unicast data frames**
- **When a LEC is transmitting, there may be two possible paths for data**
  - The unknown destination forwarding to the BUS
    - the BUS will forward the data over a P-to-MP to all LECs
  - Later, when the LE_ARP is resolved
    - the data direct VCC to the destination LEC
- **Having multiple paths introduces the possibility of out of order delivery of frames**

LANE Flush Protocol

- **LECs possible data paths**

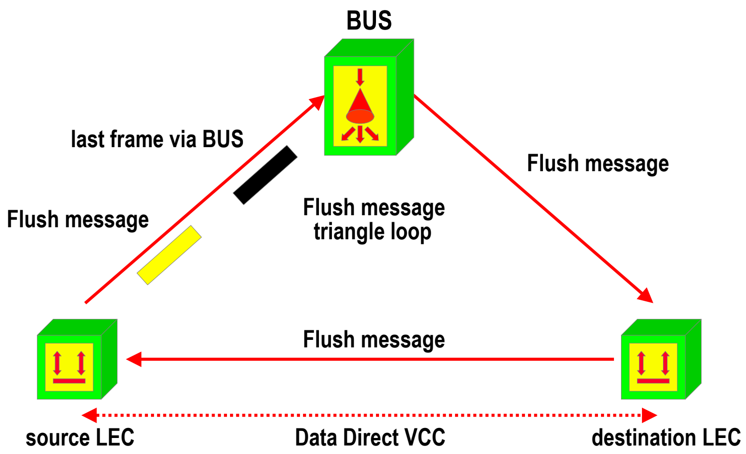


LANE Flush Protocol

- **Designed to transition traffic**
  - From the unknown unicast path (BUS) to a data direct VC
  - prevents out of order delivery of frames
- **Source LEC sends flush message to the destination**
  - along the unknown path after data direct VC has been set up
- **The LEC will hold back until the flush response message comes back, both paths are not usable**
- **Once the flush reply message is received, the new data direct path is used**

 IP over ATM, v4.2 59

Unicast Frame Synchronization



 IP over ATM, v4.2 60

LANE Concerns

- **Redundant LECS**
  - Can be implemented using existing LANE standard
  - Client needs to retrieve an additional LECS address via ILMI if call setup to first address fails
    - MAY option in LANE specification
- **Scalability - how many stations per VLAN ?**
  - Some guidelines
    - 1000 if protocol is IP
    - 500 if protocol is IPX
    - 200 if protocol is AppleTalk
- **Connectivity between ELANs**
  - Currently external device, solved with MPOA

LANE 1.0 Limitations

- **No LANE service redundancy**
  - Only one LECS, LES and BUS per ELAN
- **Limited scalability features**
  - No QOS (UBR only)
  - No ABR
  - Mo multiplexing of multiple flows over a single VC
  - No special multicast capabilities (groups)

LAN Emulation Version 2

- **Subdivided into**
  - LUNI 2.0
    - User to Network Interface
  - LNNI 2.0
    - Network to Network Interface (between servers)
- **LUNI enhancements**
  - Quality of Service (QOS) support
    - QoS for Data Direct VCCs
    - Default is UBR, QoS is optional
    - LECs may have multiple QoS sets defined that can be used by higher layers (LE_QOS_DEFINE.request)
  - Available Bit Rate (ABR) support

LAN Emulation Version 2

- **LUNI enhancements**
  - Selective Multicast Service (SMS)
    - Clients may register multicast addresses (with LES)
    - LE_ARP can be used to resolve multicast addresses
      - ATM address will always be address of BUS
    - Client may use special multicast send VC to BUS
      - Multicast traffic will only be forwarded to clients in multicast group
  - Flow multiplexing over a single VC
    - Only for data direct VCs
- **LNNI**
  - LANE server redundancy
    - Multiple LECS (one primary)
    - Redundant LES/BUS pairs (can operate in load sharing mode)

Agenda

- Introduction
- Encapsulation
- Classical IP and ARP over ATM (ARP Server)
- MTU and Signaling
- Multicast Support (MARS, MCS)
- LANE
- Next Hop Resolution Protocol (NHRP)
- MPOA

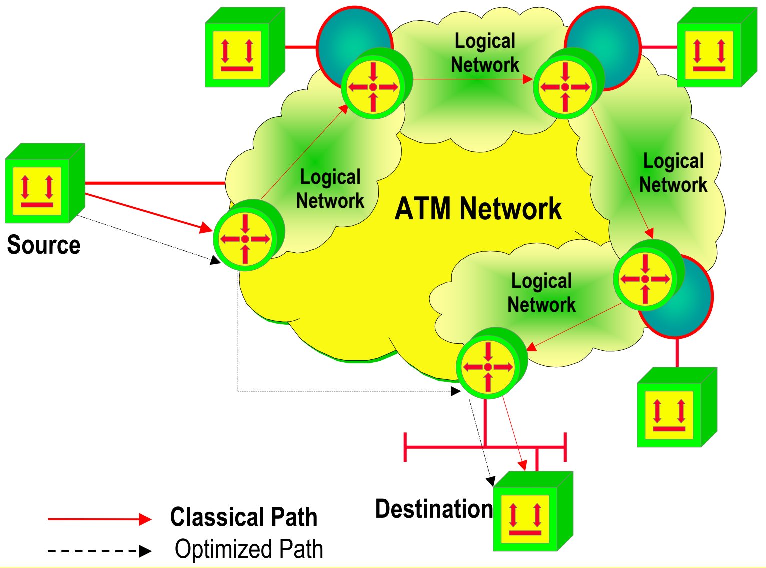
 IP over ATM, v4.2 65

Optimized Connectivity

- ATM is NBMA (Non-Broadcast Multi Access)
- Present models and protocols only suitable for one logical network
  - ATM is treated as being another LAN technology
  - Obviously a strong drawback
  - An ATM cloud will consist of several logical networks
  - Communication between logical networks
    - Always needs an external router
- Large ATM networks
  - With present models, packets have to pass several routers to reach the destination
  - However, direct link layer connection could be established

 IP over ATM, v4.2 66

Optimized Connectivity



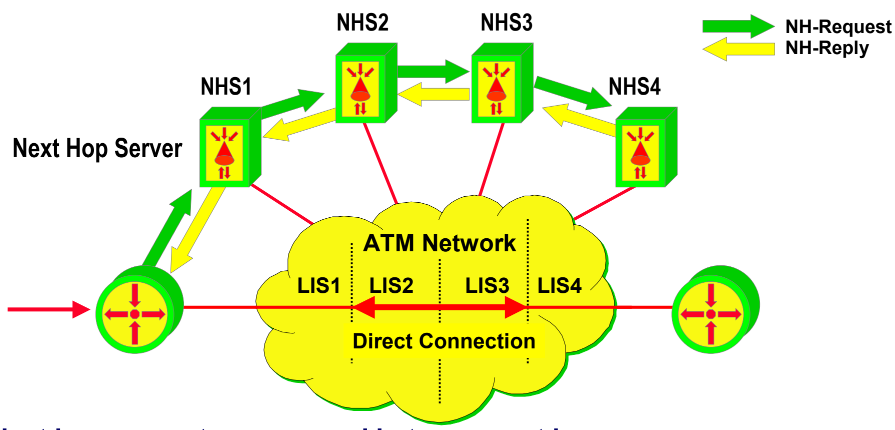
 IP over ATM, v4.2 67

Next Hop Resolution Protocol

- **Routing over Large Clouds (ROLC) working group**
  - Goal is to eliminate “extra hops” when routing over ATM
  - ATM Forum MPOA closely aligned with IETF work
- **Next Hop Resolution Protocol**
  - extension of ATM-ARP (RFC 1577)
- **NHRP takes a general approach**
  - covers end systems not connected to the same link-layer cloud
- **Is used by a source station to determine the best link layer address to reach a specific destination**
  - Destination itself / closest exit router to destination
  - Intermediate router (if policy is in effect)

 IP over ATM, v4.2 68

Next Hop Resolution Protocol (RFC 2332)



- Next hop requests are passed between next hop servers
  - Next hop servers do not forward data
- NHS that knows about the destination sends back a NH-reply
  - Allows direct connection between logical IP subnets across the ATM cloud
  - Separates data forwarding path from reachability information

Protocol Operation

- Client registers with its Next Hop Server (NHS)
  - NHRP supports multiple network layer protocols
- NHS maintains a table of address mappings
  - Derived from client registrations, address resolutions, etc.
  - All nodes using the NHS
  - Networks reachable through routers served by NHS

Protocol Operation

- **Source sends NHRP request to NHS**
  - May also send data along the (non optimized) path
- **If serving the destination, NHS will directly reply**
- **If not, request is forwarded to next NHS along the routed path**
- **Subsequent forwarding until NHS serving the destination is reached**
- **Last NHS sends reply with optimized link-layer address to source**
  - Source may now open a direct connection to this address

Agenda

- **Introduction**
- **Encapsulation**
- **Classical IP and ARP over ATM (ARP Server)**
- **MTU and Signaling**
- **Multicast Support (MARS, MCS)**
- **LANE**
- **Next Hop Resolution Protocol (NHRP)**
- **MPOA**

Overview

- **Network layer protocols use routers to communicate across subnet boundaries**
- **LANE emulates the behavior of a legacy LAN**
  - Routers are still required to communicate between ELANs
  - performance problem (SAR delay)
- **MPOA integrates LANE and NHRP**
  - Initially, traffic follows the classical model
    - Traffic is forwarded to the default gateway
    - Traffic follows the hop-by-hop routed path
  - Flow detection done by end system
  - Establishment of a short cut path to the destination

MPOA Components

- **MPOA Server (MPS)**
  - Implements the server side of the MPOA protocol
  - Includes NHRP Server (NHS)
  - Needs routing functionality, implemented in routers
- **MPOA Client (MPC)**
  - Client side of the MPOA protocol
  - Performs flow detection and requests shortcuts
- **MPOA Host**
  - Contains standard LAN Emulation Client
  - In addition, contains MPOA Client

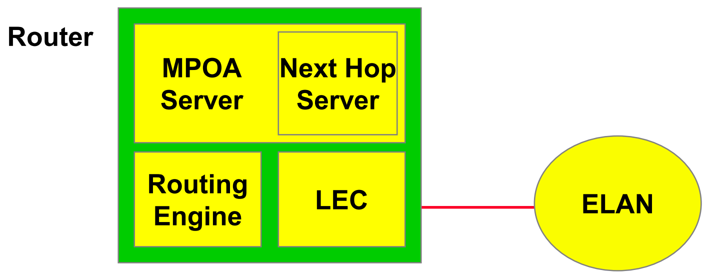
MPOA Components

- **LAN Emulation Server**
  - Identifies MPOA components
  - Clients register as MPOA capable with LES
- **LAN Emulation Configuration Server**
  - Configures MPOA clients and servers
  - Tells MPOA clients when to request a shortcut

 IP over ATM, v4.2 75

MPOA Server

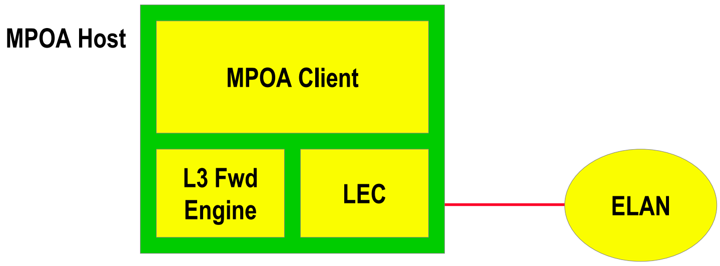
- **NHRP server co-located with a router**
  - Includes MPOA specific extensions
- **Uses NHRP to resolve IP addresses to ATM addresses**
  - Crosses over logical address boundaries
  - Follows the normal routed path across the ATM network



 IP over ATM, v4.2 76

MPOA Host

- **LANE client co-located with MPOA client**
  - Includes MPOA specific extensions
- **Does not register with the MPOA server**
  - Please note the difference to LANE
- **Identifies flows and decides when to send an MPOA request**
  - Answers to MPOA requests, called Cache Imposition Requests



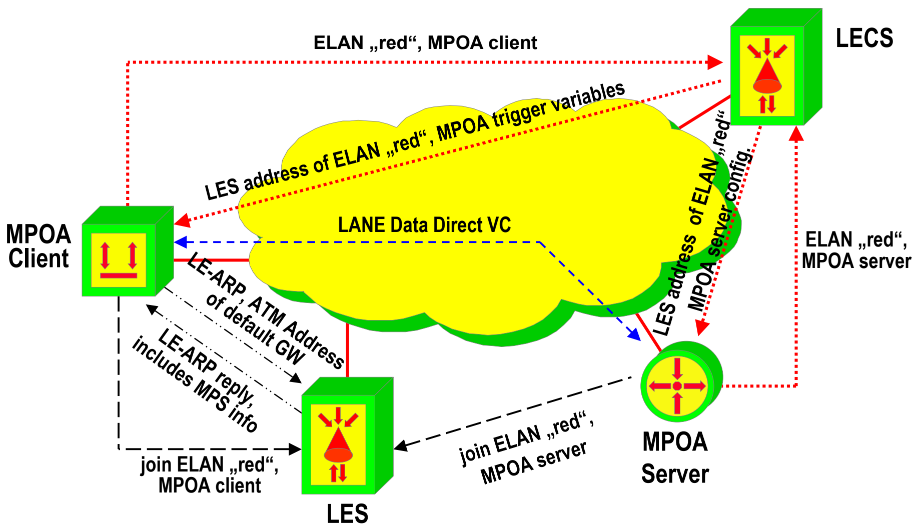
MPOA Host

- **Ingress MPOA client**
  - Flow detection
  - MPOA requests based on NHRP
  - Establishment of shortcut paths
- **Egress MPOA client**
  - Receives traffic over shortcut path
  - Adds appropriate layer 2 encapsulation and forwards the traffic to the associated LEC
  - Cache imposition requests
    - Used to inform the egress client of a shortcut

MPOA Configuration

- **Clients learn dynamically about MPOA servers**
- **MPOA clients and servers are configured by LECS**
  - Extensions to LEC configuration
  - Includes parameters for shortcut trigger values
- **Joining an ELAN**
  - Clients and servers identify themselves as being MPOA capable
  - Include MPOA TLVs (type length variables) in their registration messages

MPOA Server Discovery



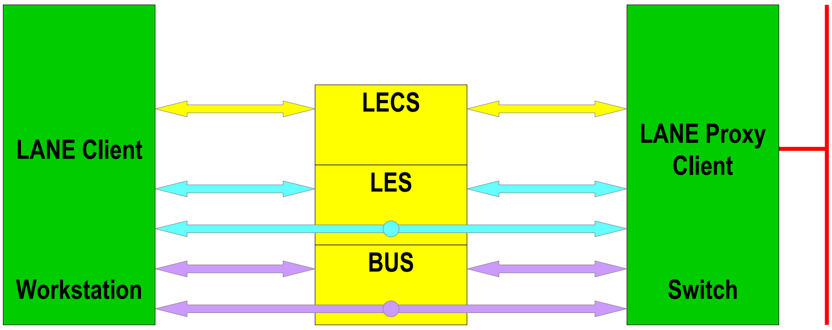
MPOA Server Discovery

- **LEC/MPC**
  - Contacts LECS and asks for information about ELAN
  - Requests how to behave in the MPOA system
- **LECS**
  - Returns LANE standard information (LES address, etc.)
  - Returns info to tell the MPC when to trigger a shortcut
- **LEC/MPS**
  - Contacts LECS and asks information about ELAN
  - Identifies itself as an MPS
- **LECS**
  - Returns LANE standard information
  - Answers with config information for MPS

MPOA Server Discovery

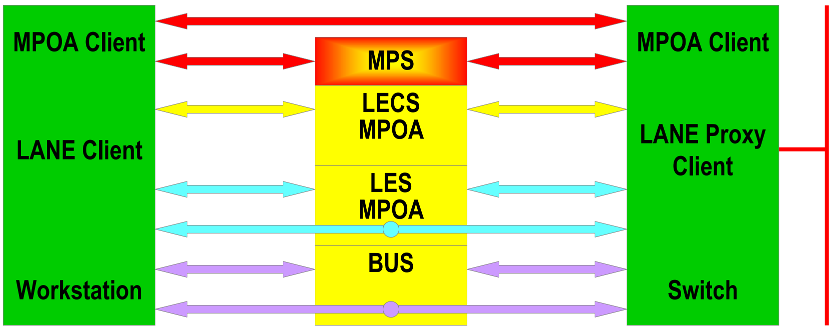
- **LEC/MPC**
  - Contacts LES and joins ELAN, identifies itself as an MPC
- **LEC/MPS**
  - Contacts LES and joins ELAN, identifies itself as an MPS
- **LEC/MPC**
  - Resolves ATM address of default gateway
  - Reply from LES contains indication that this is an MPS
  - Connects to MPS via LANE data direct circuit VC
    - note: this connection is NOT used to send MPOA requests
  - Connects to MPS via MPOA request circuit
    - note: this is the connection used for MPOA requests

Control Flow Diagram LANE



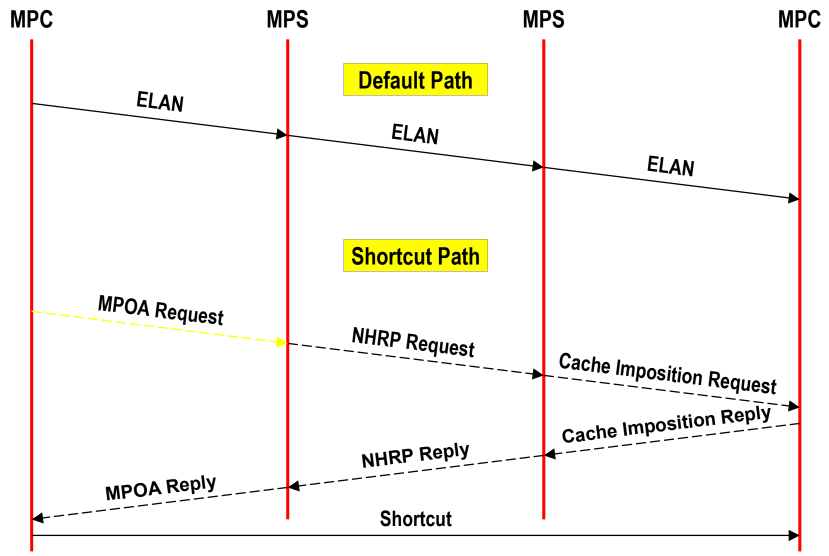
1. Configuration Direct VC between LEC and LECS
2. Control Direct VC between LEC and LES
   Control Distribute VC between LES and LECs
3. Multicast Send between LEC and BUS
   Multicast Forward between BUS and LECs

Control Flow Diagram MPOA



1. **Control Circuits between MPC and MPS**
   note: not used to forward traffic to the MPS
2. **After address has been resolved, shortcut is established**

MPOA Control Flows



 IP over ATM, v4.2 85

MPOA Operation Default

- **The MPC sends the packet in a LANE frame to the router (might be a MPS)**
  – Uses a standard LANE Data Direct VC
- **The router forwards the packet in a LANE frame to the next router (hop-by-hop forwarding) or to the destination host**
  – Via another Data Direct VC

 IP over ATM, v4.2 86

MPOA Operation — Shortcut 1

- **Ingress MPC detects a flow**
- **MPOA Resolution request sent to MPS**
  - Asks for the ATM address of a given layer 3 address
- **MPS translates request to NHRP request**
  - Sends request to next MPS
- **Last MPS checks, whether destination is an MPC**
  - Has learned that information from LEARP responses
  - Sends cache imposition request to MPC
    - Notifies the egress MPC, that someone is requesting a shortcut
    - Includes mapping between IP destination and data link header

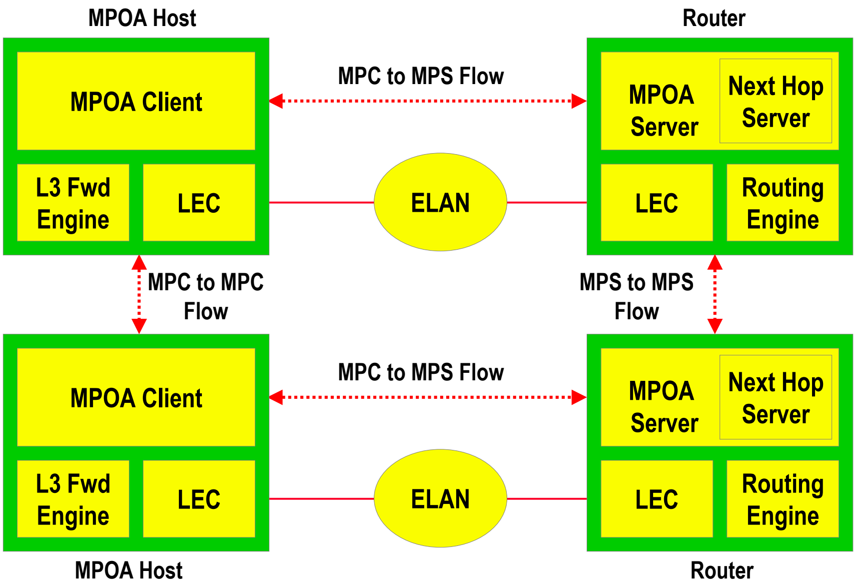
 IP over ATM, v4.2 87

MPOA Operation — Shortcut 2

- **Egress MPC receives a request**
  - Must reply, either positive or negative
  - If client has enough resources, positive reply
    - Response contains ATM address
    - Reply might contain an optional tag value
      - tag might be used as an index value for cache lookup
- **Last MPS sends back an NHRP response**
  - Either positive or negative, based on MPOA reply
- **First MPS sends MPOA response to ingress MPC**
  - Ingress MPC establishes direct VC to egress MPC
    - note: if the MPOA response is negative, the ingress MPC just continues to send data via the default path

 IP over ATM, v4.2 88

MPOA Information Flows



 IP over ATM, v4.2 89